Transferring Knowledge From One System To Another

Charles W. Bishop, Ph.D, and Peter D. Ewing, M.D.,
Department of Medicine, School of Medicine and Biomedical Sciences,
University at Buffalo, State University of New York, Buffalo, NY

Although knowledge is contained in many systems, moving it from one system to another is not an easy task because each system is tailored in its own unique way and because knowledge configurations are usually copyrighted. To populate our FRAMEMED knowledge base we turned to the NLM Metathesaurus as a readily-available open source of knowledge. We were disappointed by the greatly variable granularity of the concepts and the lack of definitions that could be borrowed. Some reference books in electronic form seem attractive but reformatting will require excessive human intervention and copyright negotiation.

INTRODUCTION

Knowledge is ubiquitous and readily accessible. Moving it from one knowledge base to another without excessive human intervention is difficult because each knowledge base is organized along different lines. Additionally, most knowledge bases are copyrighted. FRAMEMED [1] has blocked out its organization for all medical knowledge but is faced with the problem of populating its knowledge base by importing reliable medical knowledge from whatever sources it can utilize.

METHODS / RESULTS

The National Library of Medicine's Metathesaurus [2] based on equating similar concepts from multiple systems was initially attractive. Difficulty soon arose because the various systems (eg, ICD, SNOMED) were not completely hierarchical (layers not distinctively coded) and had varying degrees of granularity, relating to the purpose for which these systems were created. Thus FRAMEMED found those lists to be of limited value in constructing its hierarchical lists (as well as identifying distinct concepts about which knowledge might be sought).

An additional difficulty in trying to utilize the Metathesaurus was the plethora of files (49). The 5 unit record files were most useful but each was as large as 30 megabytes and required copying to hard drive before it could be manipulated. Examination of records truncated to eliminate formatting characters and utilizing only 6 element abbreviations revealed a preponderance of chemical entries, most of which seemed of little immediate interest. Because of their origin in schemes devised for quite specialized purposes, many entries (eg, STR0001001 Inflammatory spondylopathies in diseases classified elsewhere) were unmatchable to the hierarchical concepts in the FRAMEMED knowledge base.

We were disappointed to find that many concepts in the Metathesaurus lacked definitions and that many of the definitions present were from Dorland's Medical Dictionary and hence copyrighted.

An encouraging development for knowledge transfer is appearance of some traditional reference books such as 'The Merck Manual' in electronic form. Their definitions, because they are very context-oriented, require extensive reediting before they might contribute to the FRAMEMED descriptive records. Much material in the Merck Manual could be invaluable for our relational records of diseases (eg, disease profiles) but cannot be immediately incorporated into our system because of differing phraseology for similar concepts.

In summary, the transfer of knowledge from one system to another is not easily accomplished at present because of diversity of system design (and terminology), and copyright prohibitions. Networking may be less successful in facilitating knowlege transfer than present optimists assert, mainly because of lack of standardization of terminology and formatting.

References.

[1] Bishop CW, Ewing PD. FRAMEMED, a prototypical medical knowledge base of unusual design. M.D. Computing, 1993,10:184-92
[2] UMLS Knowledge Sources. Unified Medical Language Sytem. U.S. Dept. of Health and Human Services, National Institutes of Health, National Library of Medicine, 4th Experimental Edition, April 1993, Bethesda, MD

0195-4210/94/$5.00